# Non-linear bias with Higher Order Statistics

Luís Teodoro (LANL) Michael S. Warren (LANL)

Istvàn Szapudi (IfA, U. Hawaii) Jun Pan (IfA, U. Hawaii)

and the SDSS collaboration.

Los Alamos
NATIONAL LABORATORY

# Abstract

To decode the available data, the distribution of the underlying dark matter has to be inferred from the galaxy distribution. In principle, one can assume the two distributions are quite different, i.e., galaxies are biased tracers of the underlying density field. For a class of phenomelogical models, galaxy fluctuations $\delta_g$ are assumed to be a monotonic function of the matter fluctuation field $f(\delta_m)$.

Hence, an accurate knowledge of such a function is required to interpret galaxy statistics in light of theories of structure formation/evolution.

Los Alamos
NATIONAL LABORATORY

# Discrete Case

For the discrete probability distributions $P_N$ the generating function is defined as

$$P(x) = \sum_{N=0}^{\infty} P_N x^N$$

From $P(x)$ we can get $P_N$ by means of series expansion

$$P_N = \frac{1}{N!} \left( \frac{d}{dx} \right)^N P(x) \bigg|_{x=0}$$

Los Alamos
NATIONAL LABORATORY

# CIC (1)

Let $P_N(\ell)$ be the probability that a randomly thrown cell in the simulation contains $N$ particles, with implicit dependence on the cell size $\ell$.

$$\tilde{P}_N(\ell) = \frac{1}{C}\sum_{i=1}^{C}\delta^D(N_i = N)$$

where $C$ is the number of cells thrown and $N_i$ is the number of objects in cell $i$.

# CIC (2)

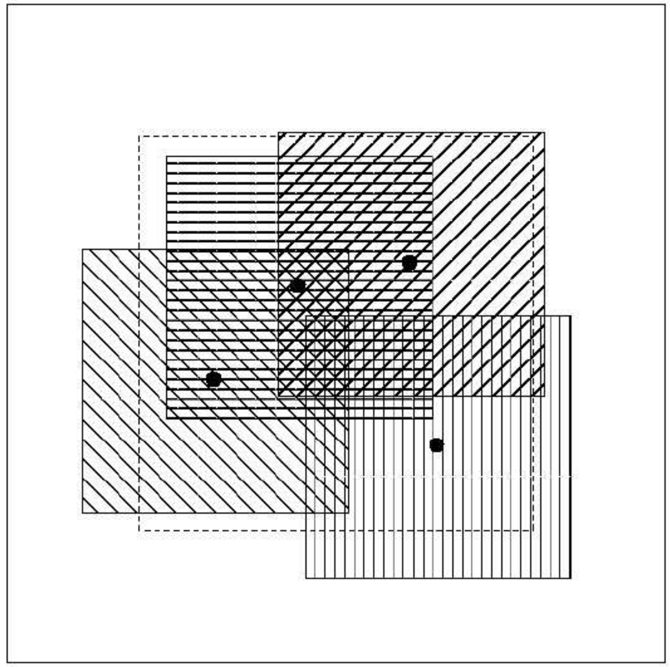
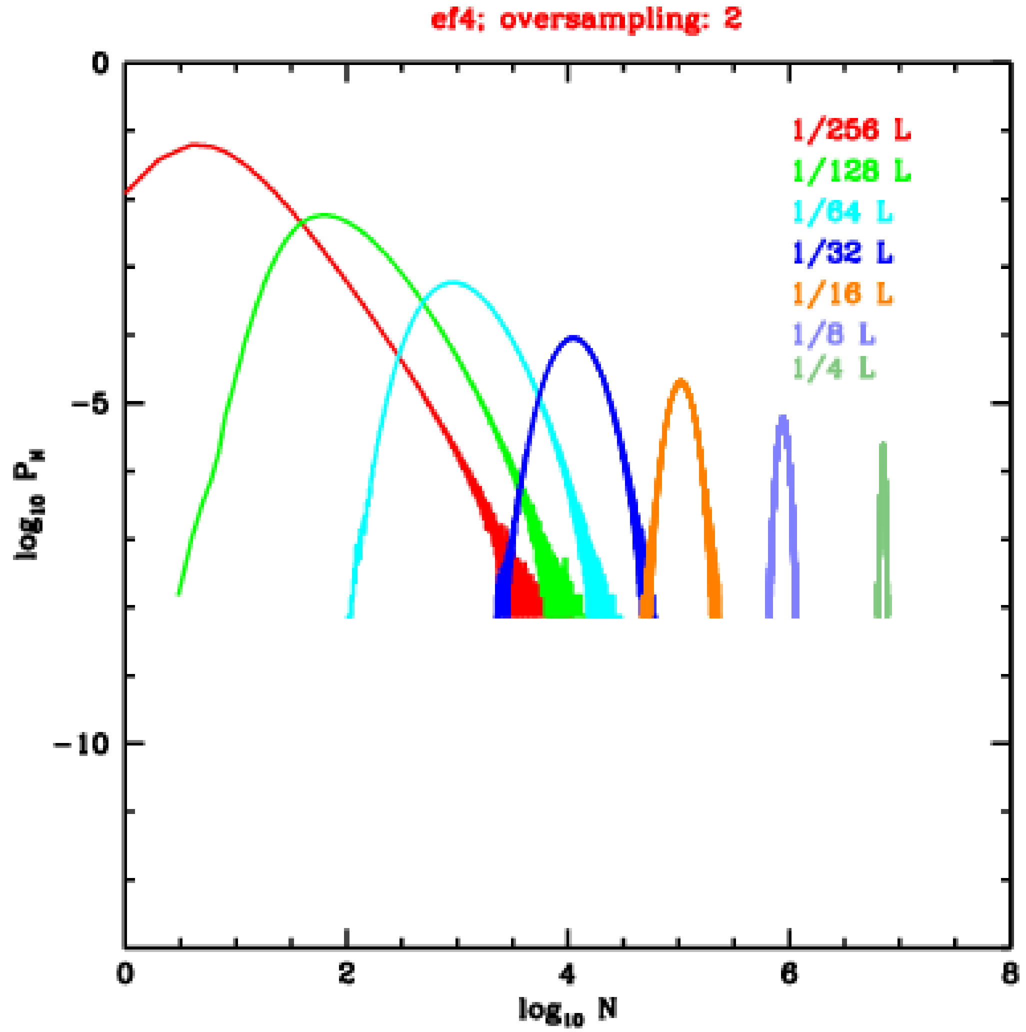


Figure 1: Illustrates the geometric calculation of counts in cells. There are four points within the solid boundary. The centres of square cells can lie within the dashed boundary. Around each point a square is drawn to represent the possible centers of cells which contain that point. Szapudi et al.(1999), Szapudi & Colombi (2004)

Los Alamos NATIONAL LABORATORY




Los Alamos
NATIONAL LABORATORY

# From Galaxies to Mass

The cumulative probability distribution functions of the galaxy and matter density fluctuation fields, $C_g(\delta_g)$ and $C_m(\delta_m)$ obey the following relation

$$C_g(\delta_g) = \int_{-1}^{\delta_g = f(\delta_m)} p(\delta_g)\, d\delta_g = \int_{-1}^{\delta_m} p(\delta_g) \left| \frac{d\delta_g}{d\delta_m} \right| d\delta_m = C_m(\delta_m)$$

To recover the bias function we need to know both cumulative distributions

$$\delta_g = f(\delta_m) = C_g^{-1}\left[C_m(\delta_m)\right]$$

Szapudi & Pan (2004)

# CIC (3)

$P_N$ is directly related to the continuous function under the locally Poissonian approximation

$$P_N(\ell) = \int_{-1}^{\infty} p(\delta) \frac{\left[\langle N \rangle (1+\delta)\right]^N e^{\langle N \rangle (1+\delta)}}{N!} d\delta = \int_{-1}^{\infty} p(\delta) K(N,\delta) d\,\delta$$

where $K(N,\delta)$ is a Poissonian kernel, $\langle N \rangle >$ is the mean CIC and $N$ is within $[0, N_{max}]$

# Richarson-Lucy Deconvolution (1)

To invert the above equation in a model independent way, we use we use the *Richarson-Lucy* (RL) method. This is an iterative method is based on Bayes's theorem. In the probabilistic spirit of this method, the functions need rescaling

$$\hat{K} = \frac{K}{\sum_N K} \qquad \text{and} \qquad \hat{p} = p \sum_N K$$

From this, a better approximation of $\hat{p}$ is obtained

$$\hat{p}_2 = \hat{p}_1 \sum_{N=0}^{N_{\text{max}}} \frac{P_N}{P_{N,1}} \hat{K}(N, \delta)$$

Los Alamos
NATIONAL LABORATORY

# Richard-Lucy Deconvolution (2)

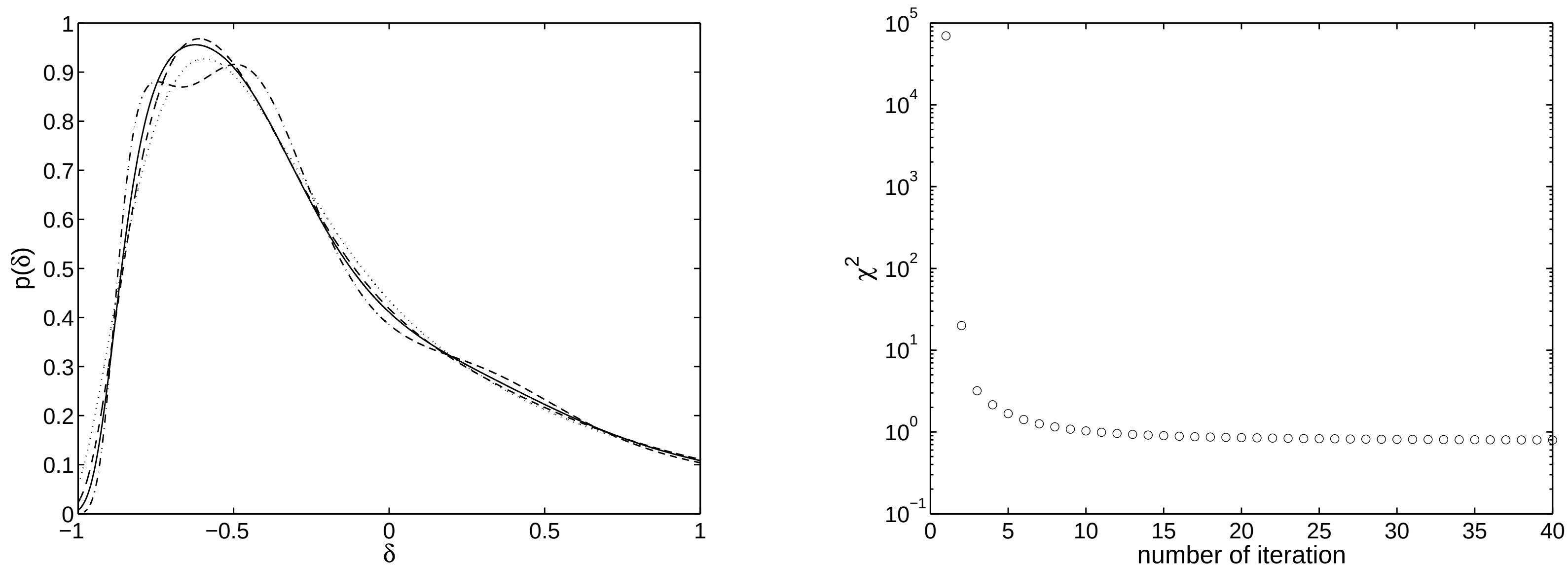
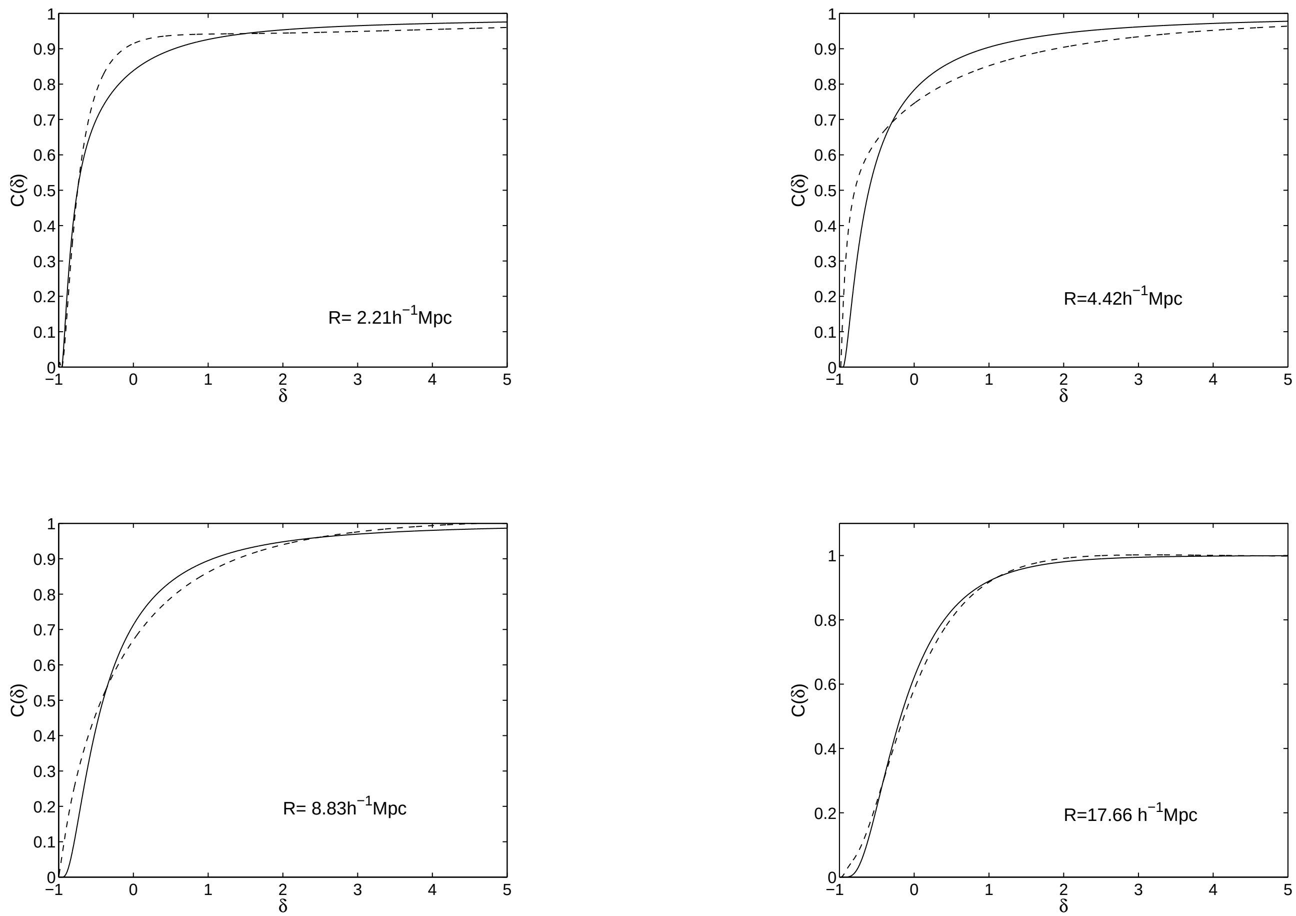



Figure 2: On the left is the RL inversion of an artificial PDF + 5% noise level for a number of different iterations [5 (dotted), 15 (dashed), 100 (dash-dotted)], while the right figure represents the cost function $\chi^2 = \sum_N (P_N/P_{N,i} - 1)^2$.

# Bias Function



# SDSS Mock CPDF

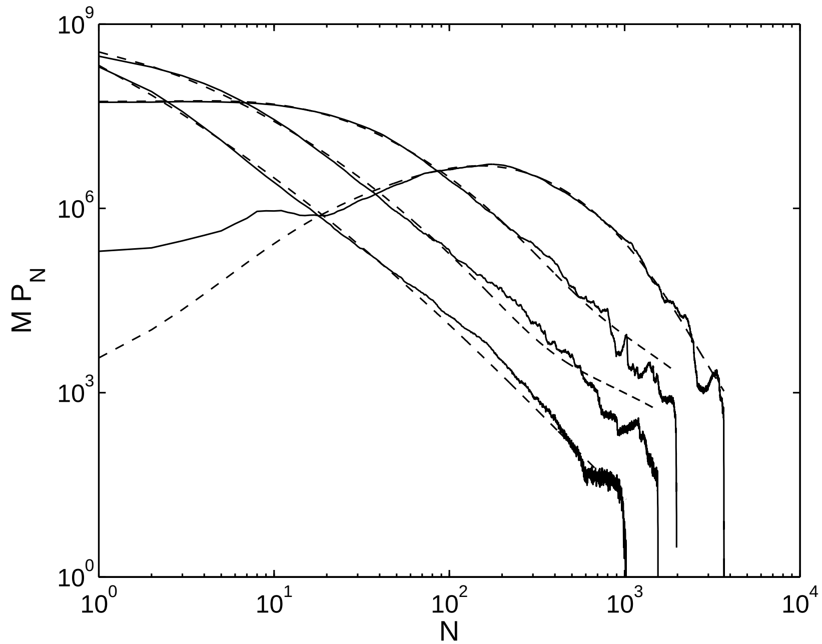
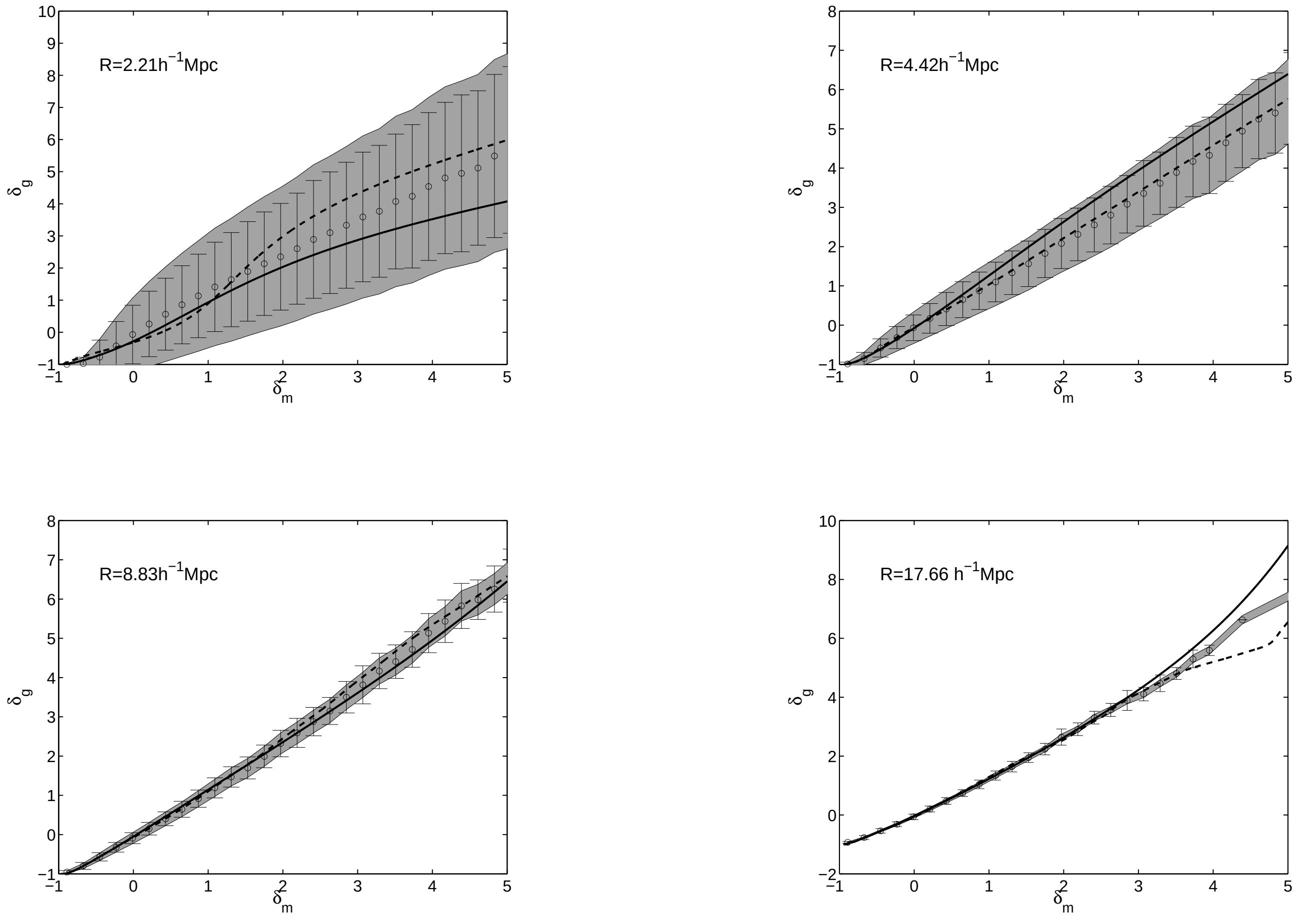



Figure 3: CPDF's SDSS mock galaxy catalog on scales of 2.21, 4.42, 8.83, and 17.66 $h^{-1}$ Mpc. GIF simulations: $L = 141.3h^{-1}$ Mpc, mass $1.4 \times 10^{10}$ $M_\odot$, $256^3$ particles "Concordance Model"

# SDSS Bias Function (GIF mocks) (1)




Los Alamos
NATIONAL LABORATORY

# Conclusions

Tests based on high-resolution dark matter simulations and corresponding mock galaxies catalogs show that we can reconstruct the nonlinear bias function down to highly nonlinear scales with precision in the range of $-1 \leq \delta \leq 5$.

Los Alamos
NATIONAL LABORATORY